Available online at www.sciencedirect.com
ScienceDirect
IFAC-PapersOnLine 48-17 (2015) 207–212

Generalised statistical process control (GSPC) in stress monitoring

Esko K. Juuso *

* *Control Engineering Group*
Faculty of Technology
90014 University of Oulu, Finland (email: Esko.Juuso@oulu.fi)

Abstract: The early detection of fluctuations in operating conditions and fault detection is done with similar methods. The feature extraction uses statistical analysis based on generalised norms and moments. Intelligent stress indices are calculated from these features by nonlinear scaling. The scaling approach uses the norms and moments to produce indices, which are consistent with the vibration severity criteria. Nonlinear scaling can be used for finding suitable control limits for the features and indices. Harmful high levels of stress are efficiently detected with control limits adjusted to the process requirements. The limits can be explained by fuzzy set systems and categorical information is included through knowledge-based analysis. The statistical process control (SPC) can be extended to nonlinear and non-Gaussian data: the new generalised SPC is suitable for a large set of statistical distributions. It operates without interruptions in short run cases and adapts to the changing process requirements. The approach is tested in two application cases: a rolling mill and an underground load haul dump (LHD) machine.

Keywords: Statisttical process control, cavitation, fatigue, vibration analysis, feature extraction, water turbine, roller mill.

1. INTRODUCTION

Strong stress levels are extremely harmful for machines and process equipment. Fatigue is caused by cracks formed when the load exceeded certain thresholds during repeated loading and unloading. Structures fracture suddenly when a crack reaches a critical size. In fatigue studies, material performance is commonly characterized by a material specific S-N curve, also known as the Wöhler curve. In practice, the sequence of load is complex, often random, including large and small loads. The effects of each stress level are taken into account in the calculations of cumulative damage from individual contributions (Palmgren, 1924; Miner, 1945). Torque measurements are informative in this analysis (Mackel and Fieweger, 2010) and can be used in fatigue prediction (Juuso and Ruusunen, 2013).

Maintenance is a critical factor in the economic performance of mining companies (Kumar, 1996; Cutifani et al., 1996). Condition-based maintenance (CBM) program could reduce the maintenance costs by removing unnecessary breakdowns. Load haul dumps (LHDs) are large loading machines that are used in the underground mining to load and move ore from the drift. The LHD machines are operated remotely. A lot of maintenance, operation and production data are available from the operation of mobile machines (Gustafson and Galar, 2012).

Intelligent methods extend the idea of dimensionless indices to nonlinear systems: the basic idea is nonlinear scaling, which was developed to extract the meanings of variables from measurement signals (Juuso, 2004). In the present systems, the scaling functions are developed by using generalised moments and norms (Juuso and Lahdelma, 2010; Juuso, 2013) and tuned with genetic algorithms (Juuso, 2009). The condition monitoring applications are similar with detecting operating conditions in the process industry (Juuso and Leiviskä, 2010). Process and condition monitoring data is combined in detecting operating conditions: measurements which require signal processing are denoted as signals, normal process measurements are directly used in feature extraction, and in addition some infrequent measurements need to be interpolated. Condition and stress indices can be obtained directly as scaled features or combinations of them. The condition indices are compared in similar load conditions but the stress indices need to be defined in non-stationary operating conditions.

Statistical process control (SPC) is based on continuously analysing and reducing variation in manufacturing processes (Oakland, 2008). Focus is on early detection and various control charts have been developed and widely used for that: Shewhart started already in1920s. Standard control charts are often based on normal distributions, but non-Gaussian data need to be analysed in many cases. A flexible family of statistical distributions is applied in (Fournier et al., 2006). In the applications, which require stress monitoring, both process measurements and condition monitoring measurements are highly nonlinear. Short run SPC introduces additional challenges for parameter estimation (Celano et al., 2013).

This paper introduces a generalised statistical process control (GSPC) for stress monitoring by using the nonlinear scaling methodology to evaluate limits. The approach is tested in two application cases: a rolling mill and an underground load haul dump (LHD) machine.

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2015.10.104

2. NONLINEAR SCALING

Meanings of feature and index levels are essential in stress monitoring. Membership functions used in fuzzy logic are represented with membership definitions, which provide nonlinear mappings from the operation area, defined with feasible ranges, to the linguistic values represented inside a real-valued interval [-2,2]. The basic scaling approach presented in (Juuso, 2004) has been improved later: a new constraint handling was introduced in (Juuso, 2009), and a new skewness based methodology was presented for signal processing in (Juuso and Lahdelma, 2010). Membership definitions are monotonously increasing scaling functions: two second order polynomials and their inverse functions.

The concept of feasible range is defined as a trapezoidal membership function. In the fuzzy set theory (Zimmermann, 1992), support and core areas are defined by variable, x_j, specific subsets,

$$supp(F_j) = \{x_j \in U_j \mid \mu_{F_j}(x_j) > 0\}, \quad (1)$$

$$core(F_j) = \{x_j \in U_j \mid \mu_{F_j}(x_j) = 1\}, \quad (2)$$

where U_j is an universal set including F_j; $\mu_{F_j}(x_j)$ is the membership value of x_j in F_j. The main area of operation is the core area, and the whole variable range is the support area. For applications, a trapezoidal function providing linear transitions between 0 and 1 is sufficient (Fig. 1). The corner parameters can be defined on the basis of expert knowledge or extracted from data. The slope can be different on upper and lower part depending on the linearity or nonlinearity of the system.

The SPC focuses on the complement, which is in (Zimmermann, 1992) defined as a subset

$$\bar{F}_j = \{x_j \in U_j \mid \mu_{\bar{F}_j}(x_j) = 1 - \mu_F(x_j)\}, \quad (3)$$

where $\mu_{\bar{F}_j}(x_j)$ is the membership value of x_j in $\bar{F}_j$. The membership function of the complement corresponds to the highest and lowest membership functions (Fig. 1).

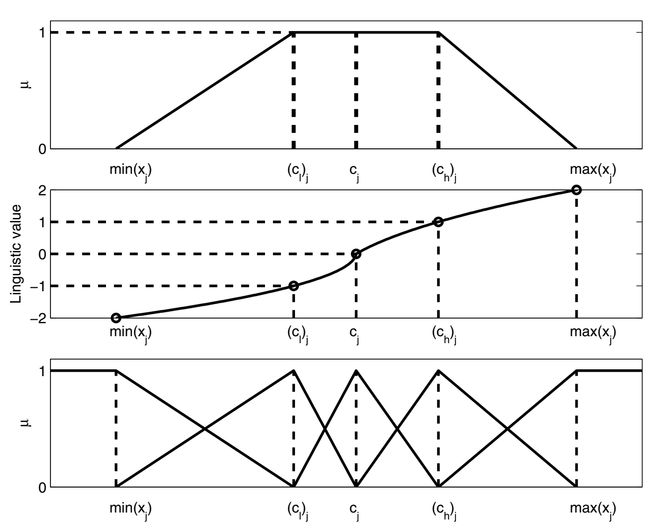

Figure 1. Feasible range, membership definitions and membership functions (Juuso, 2004)

The support area is defined by the minimum and maximum values of the variable, i.e. the support area is $[\min(x_j), \max(x_j)]$ for each variable $j, j = 1, \ldots, m$. The central tendency value, c_j, divides the support area into two parts, and the core area is defined by the central tendency values of the lower and the upper part, $(c_l)_j$ and $(c_h)_j$, correspondingly. This means that the core area of the variable j defined by $[(c_l)_j, (c_h)_j]$ is within the support area.

3. DATA ANALYSIS

The corner points can be extracted from existing rule-based fuzzy systems or defined manually. Feasible ranges should be consistent with membership definitions, and therefore they are defined together in the data-driven approach. Earlier the analysis of the corner points and the centre point has been based on the arithmetic means or medians of the corresponding data sets (Juuso, 2004).

The norm defined by

$$||^{\tau} M_j^p||_p = (^{\tau} M_j^p)^{1/p} = [\frac{1}{N} \sum_{i=1}^{N} (x_j)_i^p]^{1/p}, \quad (4)$$

where $p \neq 0$, is calculated from N values of a sample, τ is the sample time. With a real-valued order $p \in \Re$ this norm can be used as a central tendency value if $||^{\tau} M_j^p||_p \in \Re$, i.e. $x_j > 0$ when $p < 0$, and $x_j \geq 0$ when $p > 0$. The norm (10) is calculated about the origin, and it combines two trends: a strong increase caused by the power p and a decrease with the power $1/p$. Therefore, all the norms have same dimensions as x_j. The generalised norm for absolute values $|x_j|$ was introduced for signal analysis in (Lahdelma and Juuso, 2008). In stress monitoring, all the features and indices are positive.

The value range of x_j is divided into two parts by the central tendency value c_j and the core area, $[(c_l)_j, (c_h)_j]$, is limited by the central tendency values of the lower and upper part. The approach is based on the normalised moments generalised by replacing the expectation with the norm (10) as the central value:

$$\gamma_k^p = \frac{1}{N \sigma_j^k} \sum_{i=1}^{N} [(x_j)_i - ||^{\tau} M_j^p||_p]^k \quad (5)$$

where σ_j is calculated about the origin, and k is a positive integer. (Juuso and Lahdelma, 2010)

The moment (5) is used for estimating the central tendency value and the core area for the features x_j, $\alpha = 0$. The central tendency value is chosen by the point where the skewness changes from negative to positive, i.e. $\gamma_3 = 0$. Then the data set is divided into two parts: a lower part and an upper part. The same analysis is done for these two data sets. The estimates of the corner points, $(c_l)_j$ and $(c_h)_j$, are the points where the direction of the skewness changes. The iteration is performed with generalised norms. The scaling functions are monotonously increasing if the ratios

$$\alpha_j^- = \frac{(c_l)_j - \min(x_j)}{c_j - (c_l)_j}$$
$$\alpha^+ = \frac{\max(x_j) - (c_h)_j}{(c_h)_j - c_j} \quad (6)$$

are restricted to the range $[\frac{1}{3}, 3]$. The corner points $(c_l)_j$ and $(c_h)_j$ or the upper and lower limits $\min(x_j)$ and/or $\max(x_j)$ are moved if these requirements are violated.

The centre point is not known if the feasible range is defined manually. It can be calculated by defuzzifying the feasible range with the centre of gravity:

$$c_j = \frac{1}{4}\,((c_l)_j + (c_h)_j + \min(x_j) + \max(x_j)). \quad (7)$$

For strongly asymmetrical feasible ranges, this value may be outside the core (Juuso, 2004). The requirement (7) can be fulfilled by modifying the corner points.

Additional constraints can be taken into account, e.g. a good solution can be to use a locally linear function in the neighbourhood of the centre point. Then a continuous derivative is chosen at the centre point: $b_j^- = b_j^+$, which can be represented by

$$6\,c_j - 4\,(c_l)_j - 4\,(c_h)_j + \min(x_j) + \max(x_j) = 0. \quad (8)$$

This can be achieved by modifying the centre point or the corner points of the feasible range. There can be several acceptable modifications, for which the ratios (6) remain in the range $[\frac{1}{3}, 3]$.

Monotonously increasing membership definitions can be constructed by adjusting the centre point c_j, the core $[(c_l)_j, (c_h)_j]$ and the support $[\min(x_j), \max(x_j)]$. An easier way for manual approach was introduced in (Juuso, 2009): first define the centre point c_j, then the core by choosing the ratios (6) from the range $[\frac{1}{3}, 3]$, and finally calculate the support $[\min(x_j), \max(x_j)]$. The norms (10) are used together with the generalised skewness (5) in the data-driven approach to define the centre and corner points. The ratios (6), which are checked in all data-driven cases, are also guiding the manual construction of the membership definitions. Additional constraints like (7) and (8) are used if they are feasible.

The nonlinear scaling methodology provides good results for the automatic generation of scaling functions. Even small faults and anomalies are detected. The approach has been tested with normal, Poisson and Weibull distributions and used in condition monitoring applications (Juuso and Lahdelma, 2010). This approach is suitable for a very large set of statistical distributions (Juuso, 2013).

4. CONTROL CHARTS

Statistical process control (SPC) provides algorithms for detecting deviations from the defined operating areas of individual variables. The SixSigma approach is based on normal distributions, and therefore, it can be understood as a symmetric special case, where the shape ratios $\alpha_j^- = \alpha_j^+ = 3$. The upper and lower control limits, $[c_j + 3\sigma_j$ and $c_j - 3\sigma_j]$, correspond to the scaled values $X_j = \sqrt{6}$ and $X_j = -\sqrt{6}$, respectively. The upper and lower warning limits are $c_j + 2\sigma_j$ and $c_j - 2\sigma_j$, i.e. the scaled values $X_j = 2$ and $X_j = -2$.

The nonlinear scaling with the second order polynomials to $[-2, 2]$ provides a basis for variable specific control charts, which take into account the distributions. First the center line c_j is defined with an appropriate generalised norm. Then the shape ratios α_j^+ and α_j^- are used to define the upper and lower control limits by using the limits $(X_j)_{max}$ and $(X_j)_{min}$, and the upper and lower warning limits by using the limits of the support area, see Figure 2. The limits can be highly asymmetrical, when $\Delta c_j^+ \neq \Delta c_j^-$ and/or $\alpha_j^+ \neq \alpha_j^-$, e.g. in Poisson distributions $\alpha_j^+ = 3$ and $\alpha_j^- = 1$ with strongly asymmetrical core areas. The analysis with the scaled values X_j is beneficial in short term SPC, e.g. specific scaling functions and limits can be developed for the operating conditions.

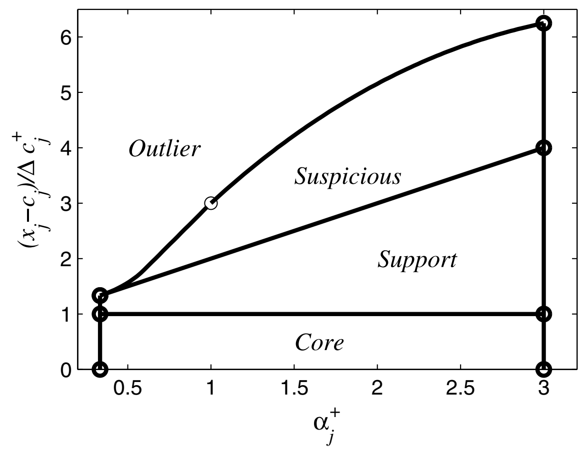

Figure 2. Limits for the core, support, suspicious and outlier areas as a function of the shape ratio α_j^+ (Juuso, 2013).

Cavitation effects on the stress of the water turbine: a quantitative measure is a cavitation index, which provides an indication of both clear cavitation and clearly good operation. Values -2 and -1 indicate good operating conditions. Value 1 corresponds to the clear signs of cavitation and value 2 means a very strong indication of cavitation. The cavitation index is obtained from vibration measurement by using derivation, feature extraction and nonlinear scaling presented above.

The generalised SPC expands the SPC from Gaussian to non-Gaussian data sets. The analysis methods are suitable for a large set of statistical distributions. Categorical information can be studied with the same approach by using manual definitions, which means that also mixed cases can be handled. The limits can be updated in short run SPC since they are defined by the nonlinear scaling approach. The limits can even change gradually. The GSPC does not need any interruptions and even recursive approaches are possible.

5. APPLICATION CASES

In application cases, the upper control limits are obtained by analysing the statistical distributions of the features related to the stress.

5.1 Roller mill

Torque measurements collected from a rolling mill have been used in the testing of the approach. The derivation is not used for these signals. The feature is a combination of two norms and the stress index is calculated from two scaled features obtained by using the nonlinear scaling approach. The resulting linguistic S-N curve is linear and a normal S-N curve is formed from it by scaling to the feature values: a large number of passes have low stress indices. The high stress cases are seen as a very steep rise in the semilogarithmic curve. At the risk level higher than 60%, a single high torque level can have a strong effect on the activation of a failure. Long operating periods can be achieved if the risk levels are low (Fig. 3). The approach operates well for the limited set of failures analysed in this study and is promising for practical use.

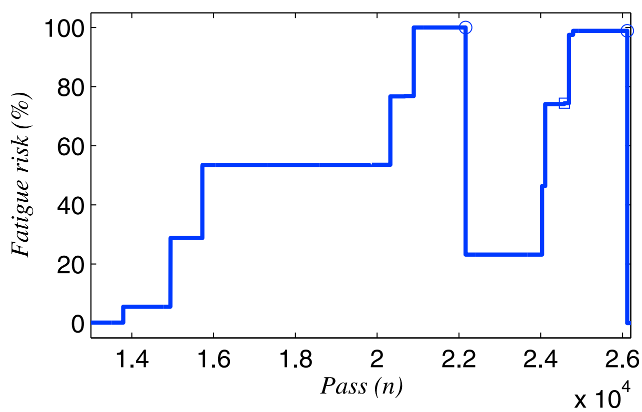

Figure 3. Calculated fatigue risk (%): o is a failure point and □ a pass with high torque, which does not cause a failured, extracted from (Juuso and Ruusunen, 2013).

Intelligent stress indices based on nonlinear scaling provide good indicators of the severity of the load. The stress index I_s is -2 when the stress is negligible, and levels $\{-1, 0, 1\}$ are analogue to the lower limits of the vibration severity ranges {usable, still acceptable, not acceptable}. The Wöhler curve is represented by a linguistic equation (LE) model

$$I_S = log_{10}(N_C) \quad (9)$$

where the stress index can be a scaled value of stress or a scaled value of a generalised norm obtained from signals (Juuso and Ruusunen, 2013). The contribution of the stress is calculated in each sample time τ, which is taken as a fraction of the cycle time N_C. The cumulative sum of the contributions indicates the deterioration of condition and the simulated sums can be used to predict failure time. The previous history can be updated whenever the scaling functions are changed.

The stress index is calculated from two norms:

$$I_S = [\frac{1}{N}\sum_{i=1}^{N}(x_j)_i^2]^{1/2} - \frac{1}{N}\sum_{i=1}^{N} x_j \quad (10)$$

obtained from torque measurements. The derivation is not used in this case. The scaling function is highly nonlinear (Fig. 4). The support area is changed to ensure that the scaling function is monotonously increasing. All the feature values are positive: the negative value corresponding to the very low level is needed for function definitions. This can be done since only the upper limits are interesting in the stress analysis.

The upper limit of the support area is modified as there are very few high values: the upper limit of support area is lowered, i.e. the values above UCL were considered being outside the support area when the scaling function was adjusted. Thus a suitable control limit was obtained by the data analysis.

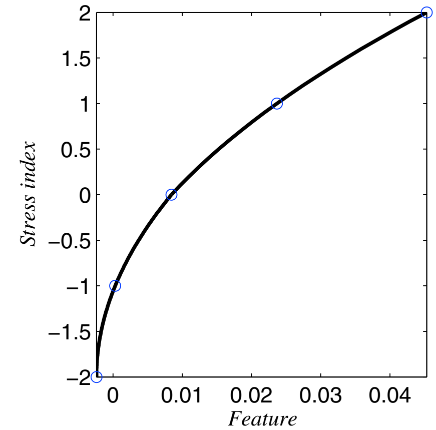

Figure 4. Scaling function of the feature.

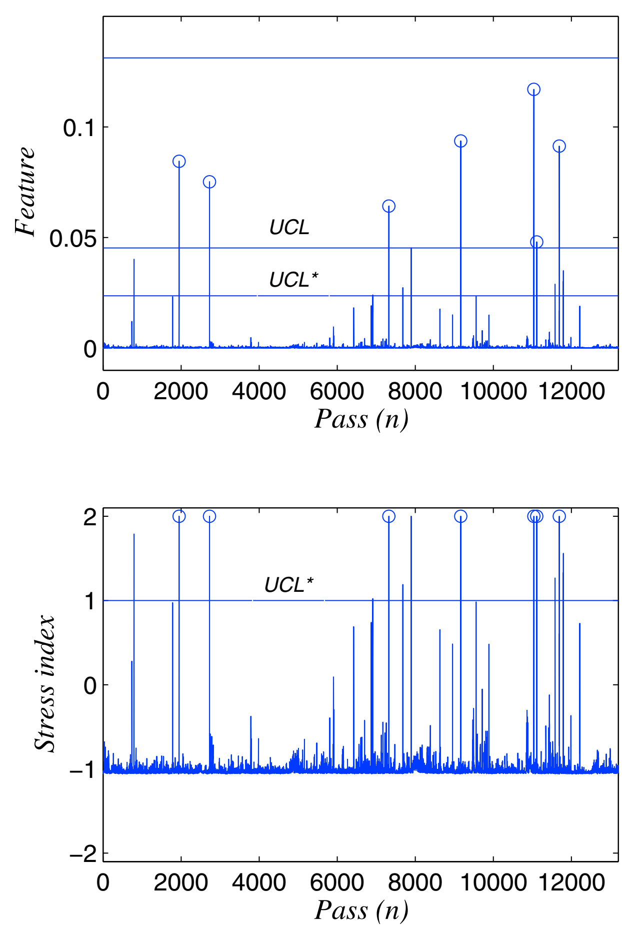

Figure 5. GSPC results of the roller mill data: $UCL = 0.0453$.

The upper control limit is defined by the level 2 of the stress index (Fig. 5). Most of the time, the stress levels are low, but the high feature levels exceeding UCL have a strong effect on the fatigue risk (Fig. 3). Several high values

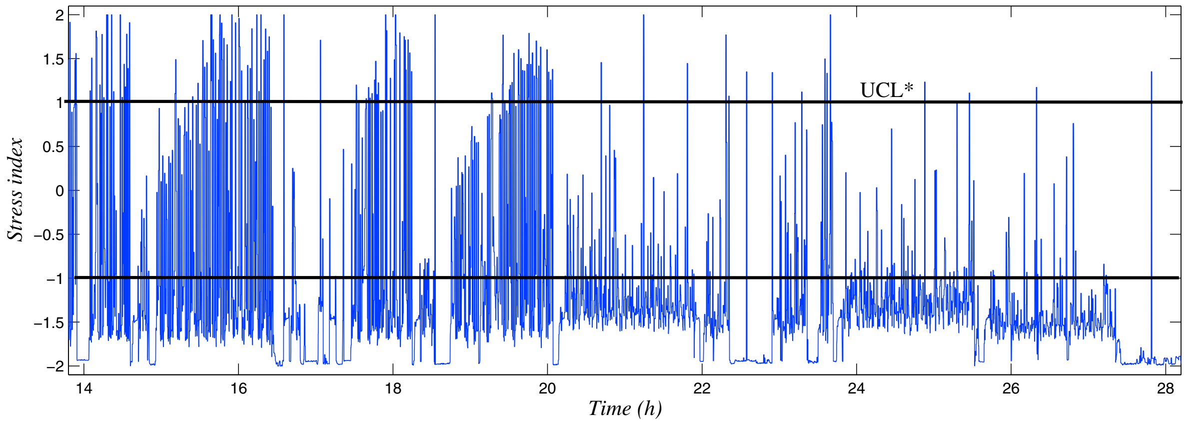

Figure 6. A short period of the GSPC results in the LHD data.

are needed to cause a failure: first two result increase in the risk, but two additional ones are required for the failure. Later three consequent high stress values are followed by a failure.

The control limit UCL corresponding to $I_S = 2$ is related to mechanical failures. Again better quality performance can be achieved if the limit is moved to UCL^*, i.e. $I_S = 1$. Removing the values exceeding the level corresponding to $I_S = 2$ will then change distribution to the quality control which is main area of the SPC.

5.2 Front axle of a LHD

The vibration measurements from the front axle of a load haul dumper (LHD) have been analysed with G. These machines operate in harsh conditions where failures may be difficult to repair. The machine under study is a Sandvik LH 261 working underground in the Pyhsalmi mine. Four accelerometers were mounted on its front axle housing. Four SKF CMPT 2310 accelerometers were mounted externally onto the LHDs front axle housing to measure horizontal and vertical vibrations near the planetary gearboxes on either side. These four vibration measurements together with a tachometer pulse from the drive shaft are recorded with a National Instruments CompactRIO 9024 data logger into a solid-state drive (SSD) as files of one minute length.

Signals were recorded with sampling frequency 12800 Hz, and a built-in antialising filter guarantees that there are no aliases at frequencies that are less than 0.45 x 12800 Hz = 5760 Hz. More information on the measurements can be found from (Laukka et al., 2015). The measurement points are right vertical (RV), left vertical (LV), right horizontal (RH) and left horizontal (LH). The measurements cover a period of 271 days and the data collection still continues. In this paper, the analyses uses only 26 days to demonstrate the possibilities of detecting stress levels with vibration measurements. The scaling functions are defined from the data of this period.

A seven day period shown in Figure 6 demonstrates the differences: three high stress days are followed by four days of low stress. The GSPC operates in the same way as in the roller mill case. The control limit UCL corresponding to $I_S = 2$ is related to harmful effects. Again better quality performance can be achieved if the limit is moved to UCL^*, i.e. $I_S = 1$. Several high values are needed to cause a failure. Cumulative stress increases during the high risk periods and there are considerable differences between the measurement points (Fig. 7).): value -2 means no stress, values below zero are considered negligible stress; only stress indices > 1 are taken into account in the cumulative stress. The levels {-1, 0, 1} are analogue to the lower limits of the vibration severity ranges {usable, still acceptable, not acceptable} defined in (VDI, 1964; Collacott, 1977). Maintenance actions were needed shortly after this measurement period.

Removing the values exceeding the level corresponding to $I_S = 2$ will then move the emphasis to the quality control which is main area of the SPC. The generalised SPC, which can be adapted to all the measurement points, provides useful information for the operator in the remote control of the LHD.

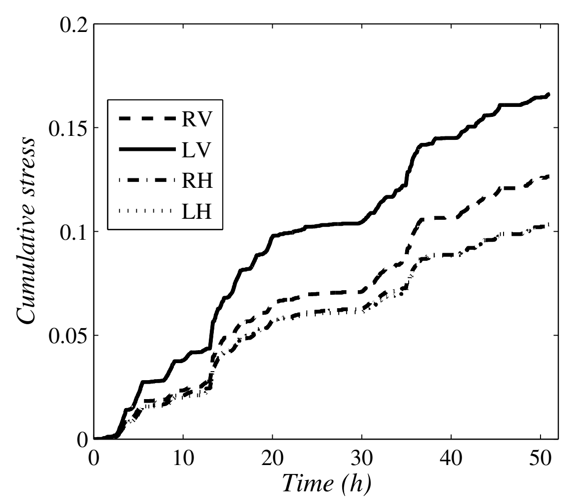

Figure 7. Cumulative stress index calculated for 26 days from two hour daily measurement periods: four measurement points (Juuso, 2014).

6. CONCLUSIONS

The early detection of fluctuations in operating conditions and fault detection can both be based on generalised norms and moments. The new scaling approach, which also uses the norms and moments, improves sensitivity to small fluctuations. The overall procedure includes the following steps: (1) select informative features, (2) scale the features, (3) calculate intelligent indices, and (4) combine indices in models. As the analysis is based the same methodology in all these applications, monitoring of the machines can be combined with process data. The smooth operation and high quality of products is the main goal of all these applications. The generalised statistical process control (GSPC) operates here as an early detection solution whose parameters are adjusted to the process requirements.

ACKNOWLEDGEMENTS

Experiments were carried out within two projects "Integrated condition-based control and maintenance (ICB-COM)" and "Production integrated condition-based maintenance model for mining industry (DEVICO)". The funding of the TEKES (the Finnish Funding Agency for Technology and Innovation), European Regional Development Fund, Outokumpu Stainless Ltd, Sandvik Mining and Construction Ltd and First Quantum Minerals Ltd Pyhäsalmi Mine are acknowledged.

REFERENCES

Celano, G., Castagliola, P., Fichera, S., and Nenes, G. (2013). Performance of t control charts in short runs with unknown shift sizes. *Computers & Industrial Engineering*, 64, 56–68.

Collacott, R.A. (1977). *Mechanical Fault Diagnosis and Condition Monitoring*. Chapman and Hall, London.

Cutifani, M., Quinn, B., and Gurgenci, H. (1996). Increased equipment reliability, safety and availability without necessarily increasing the cost of maintenance. In *Mining Technology Conference. Freemantle WA. 10-11 September*. Cooperative Research Center for Mining Technology.

Fournier, B., Rupin, N., Bigerelle, M., Najjar, D., and Iost, A. (2006). Application of the generalized lambda distributions in a statistical process control methodology. *Journal of Process Control*, 16, 1087–1098.

Gustafson, A. and Galar, D. (2012). Fusion of production, operation and maintenance data for underground mobile equipment. In *Proceedings of 9th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, 12-14 June, 2012, London, UK*, volume 1, 783–791. Curran Associates, NY, USA. ISBN 978-1-62276-433-4.

Juuso, E. (2013). *Integration of intelligent systems in development of smart adaptive systems: linguistic equation approach*. Ph.D. thesis, University of Oulu. 258 pp., http://urn.fi/urn:isbn:9789526202891.

Juuso, E. and Lahdelma, S. (2010). Intelligent scaling of features in fault diagnosis. In *7th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies 2010, CM 2010/MFPT 2010*, volume 2, 1358–1372. URL www.scopus.com.

Juuso, E. and Leiviskä, K. (2010). Combining process and condition monitoring data. In *Maintenance, Condition Monitoring and Diagnostics, Proceedings of the 3rd International Seminar, 29-30 September, Oulu, Finland*, 120–135. POHTO, Oulu, Finland. ISBN 978-951-98113-5-2.

Juuso, E. and Ruusunen, M. (2013). Fatigue prediction with intelligent stress indices based on torque measurements in a rolling mill. In *10th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies 2013, CM 2013 and MFPT 2013*, volume 1, 460–471. URL www.scopus.com.

Juuso, E.K. (2004). Integration of intelligent systems in development of smart adaptive systems. *International Journal of Approximate Reasoning*, 35, 307–337. Doi: 10.1016/j.ijar.2003.08.008.

Juuso, E.K. (2009). Tuning of large-scale linguistic equation (LE) models with genetic algorithms. In M. Kolehmainen (ed.), *Revised selected papers of the International Conference on Adaptive and Natural Computing Algorithms - ICANNGA 2009, Kuopio, Finland, Lecture Notes in Computer Science*, volume LNCS 5495, 161–170. Springer-Verlag, Heidelberg. Doi: 10.1007/978-3-642-04921-7_17.

Juuso, E.K. (2014). Intelligent indices for online monitoring of stress and condition. In *11th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies 2014, CM 2014 and MFPT 2014*, volume 1, 637–648. URL www.scopus.com.

Kumar, U. (1996). Maintenance strategy for mechanized and automated mining systems. In Hennies, A. da Silva, and Chaves (eds.), *Mine Planning and Equipment Selection*, 553–558. Balkema, Rotterdam.

Lahdelma, S. and Juuso, E. (2008). Signal processing in vibration analysis. In *Proceedings of the Fifth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, 15-18 July 2008, Edinburgh, UK*, 867–878. BINDT.

Laukka, A., Saari, J., Ruuska, J., Juuso, E., and Lahdelma, S. (2015). Condition based monitoring for underground mobile machines. *Int. J. Industrial and Systems Engineering*. In press.

Mackel, J. and Fieweger, M. (2010). Condition monitoring in steel industry. In *Proceedings of Maintenance, Condition Monitoring and Diagnostics*, 26–47. POHTO Publications, Oulu, Finland.

Miner, M.A. (1945). Cumulative damage in fatigue. *ASME Journal of Applied Mechanics*, 67, 159–164.

Oakland, J.S. (2008). *Statistical Process Control*. Routledge, New York, 6 edition.

Palmgren, A. (1924). Die Lebensdauer von Kugellagern. *Verfahrenstechnik*, 68, 339–341.

VDI (1964). VDI 2056 Beurteilungßstäbe für mechanische Schwingungen von Maschinen - VDI-Richtlinien.

Zimmermann, H.J. (1992). *Fuzzy set theory and its applications*. Kluwer Academic Publishers.